e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

SIGN- LANGUAGE -TO-TEXTSPEECH

Tanishka Kailas Aher*1, Kashish Ramesh Ugale*2, Diksha Sunil Gaikwad*3, Vaishnavi Balu Ugale*4, Prof. K.V Karad*5

*1,2,3,4Loknete Gopinathji Munde Institute Of Engineering Education & Research, Nashik, India.

*5Professor, Loknete Gopinathji Munde Institute Of Engineering Education & Research, Nashik, India.

DOI: https://www.doi.org/10.56726/IRJMETS71813

ABSTRACT

The proposed solution is a cutting-edge real-time Sign Language to text and speech translation application designed to bridge communication gaps between the deaf and hard-of-hearing community. This application aims to enhance accessibility and inclusivity by converting sign language gestures into accurate, readable text and natural-sounding speech in multiple Indian languages, including Hindi, Marathi. Utilizing advanced computer vision and machine learning technologies, the application will recognize and interpret a comprehensive library of signs and gestures with high precision. Through the integration of convolutional neural networks (CNNs) and natural language processing (NLP), the system will deliver real-time translation, ensuring that users receive immediate and contextually relevant text and speech outputs. The application will feature an intuitive, user-friendly interface that simplifies interaction, making it accessible for both SL users and hearing individuals.

Additionally, adaptive learning mechanisms will continuously improve the system's accuracy based on user feedback and interactions. This innovative solution is designed to significantly empower individuals who rely on SL, fostering greater understanding and engagement in various aspects of daily life and promoting a more inclusive society.

Keywords: Sign Language, Real-Time Translation, Text And Speech Conversion, Computer Vision, Machine Learning, Convolutional Neural Networks, Natural Language Processing, Inclusivity, Accessibility, Adaptive Learning.

I. INTRODUCTION

Effective communication is fundamental to social interaction, yet individuals who use Sign Language often encounter barriers when interacting with those who do not understand their language. To address this challenge, the proposed solution is a cutting edge real-time Sign Language to text and speech translation application.

This innovative application is designed to bridge communication gaps between the deaf and hard-of hearing community and the hearing world, enhancing both accessibility and inclusivity.

The application aims to transform SL gestures into accurate, readable text and natural sounding speech in multiple Indian languages, including Hindi and Marathi. By leveraging advanced computer vision and machine learning technologies, the system will recognize and interpret a comprehensive library of SL signs and gestures with high precision.

Through the integration of convolutional neural networks (CNNs) and natural language processing (NLP), it will deliver real-time translation, ensuring users receive immediate and contextually relevant text and speech outputs.With an intuitive, user-friendly interface, the application will facilitate easy interaction for both ISL users and hearing individuals. Its adaptive learning mechanisms will continually enhance accuracy based on user feedback and interactions. This solution is designed to significantly empower individuals who rely on SL, fostering greater understanding and engagement in various aspects of daily life and contributing to a more inclusive society.

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

II. METHODOLOGY

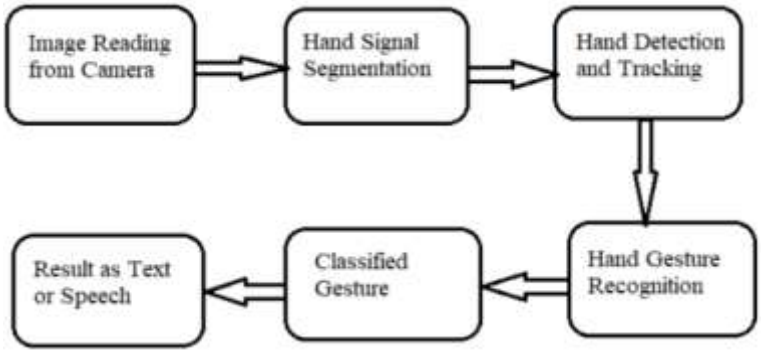

Figure 1: Block Diagram

The proposed system leverages deep learning and computer vision techniques to provide a more accurate, real-time, and scalable solution. The system aims to enhance sign language recognition by incorporating the following key components:

- Real-Time Gesture Recognition: Utilizing convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to recognize dynamic and static hand gestures efficiently.
- Multimodal Input Processing: Integrating depth sensors and RGB cameras to improve gesture detection accuracy across different lighting conditions and backgrounds.
- Speech and Text Output: Converting recognized signs into spoken language and text, enabling seamless communication between individuals with hearing impairments and the general population.
- Personalized Training and Adaptation: Allowing users to train the system on personalized gestures, improving recognition accuracy for unique or region specific sign variations.

This system aims to provide an efficient, real-time, and inclusive solution for sign language translation, bridging communication gaps and fostering greater accessibility for individuals with hearing impairments.

III. MODELING AND ANALYSIS

1. Gesture Input Module
- Users can provide input using a webcam.
- The system captures real-time video frames to detect hand gestures.
- Uses OpenCV and MediaPipe for hand tracking and segmentation.

2. Preprocessing Module
- Converts image/video frames into grayscale or binary format for better feature extraction.
- Normalizes and resizes images for consistent input size.
- Applies edge detection and contour analysis for accurate gesture identification.

3. Sign Language Recognition Module
- Utilizes a pre-trained Machine Learning (ML) model (e.g., CNN, LSTM, or Transformer-based models) for gesture classification.
- Predicts the corresponding letter, word, or phrase based on the detected hand gesture.
- Displays the recognized sign on the web interface.

4. Text Output Module
- Converts recognized gestures into readable text.
- Displays the output in real-time on the web interface.
- Provides an option to copy the recognized text.

5. Audio Output Module
- Converts the recognized text into speech using Text-to-Speech (TTS) API.
- Helps non-sign language users understand the communication.

6. UI/UX Interface Module

e-ISSN: 2582-5208
International Research Journal of Modernization in Engineering Technology and Science
(Peer-Reviewed, Open Access, Fully Refereed International Journal)
Volume:07/Issue:04/April-2025 Impact Factor- 8.187 www.irjmets.com

- A simple and interactive web-based interface using python.
- Users can start/stop the camera feed and see real-time sign detection results.
- Includes accessibility features such as dark mode and larger fonts.

7. Performance Monitoring Module

- Tracks frame processing speed and recognition accuracy.
- Displays a confidence score for each predicted gesture.
- Allows users to report incorrect predictions for model improvement.

IV. RESULTS AND DISCUSSION

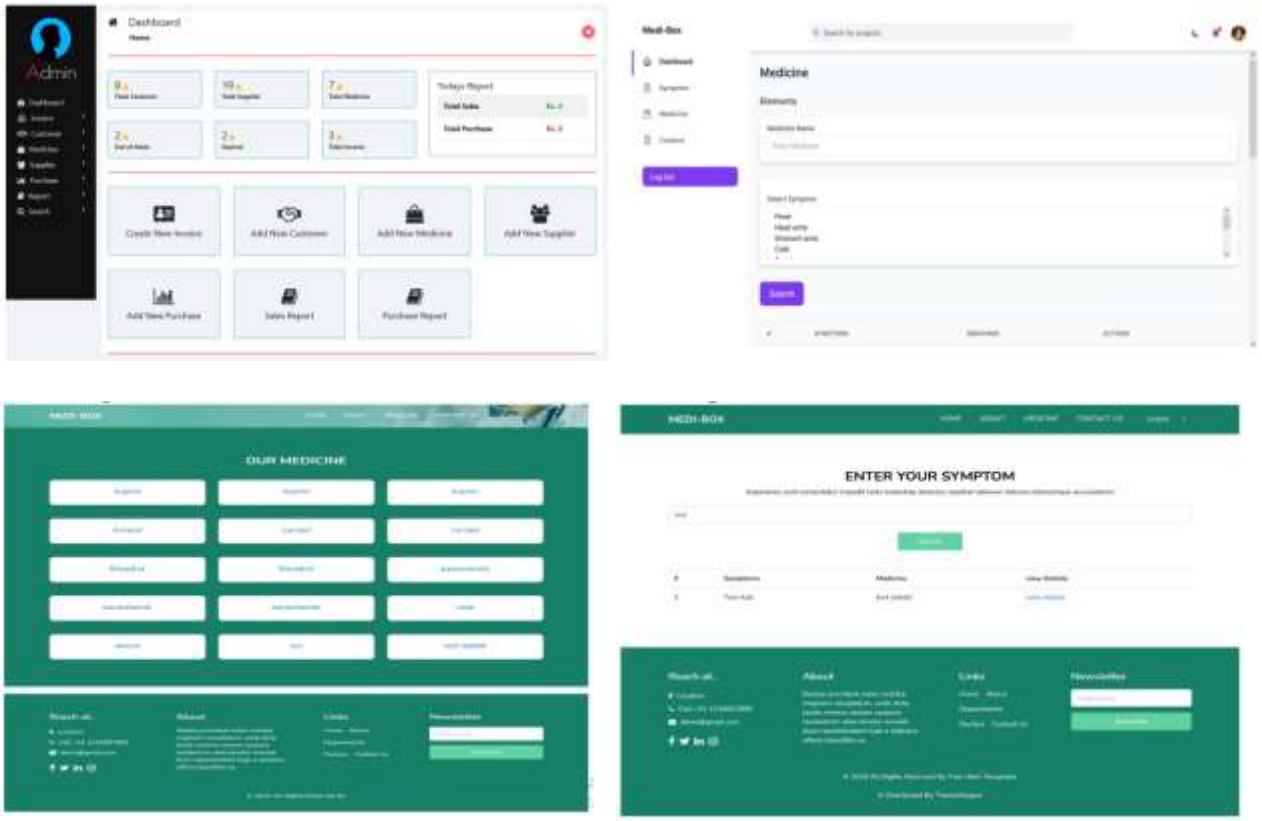

Figure 2: Results of System

A website for a medical service called "MEDI-BOX". The website is simple and clean, using a teal and white color scheme. The main focus is a search bar where users can enter their symptoms. Below that, a table displays sample symptoms and corresponding medications (though these appear to be placeholder text: "Test Add" and "test adddd"). The dashboard is well-organized and informative, designed to quickly provide an overview of key metrics and facilitate essential data entry functions. The use of upward-trending indicators suggests the system is tracking progress, possibly sales or inventory growth. The zero values in "Today's Report" suggests the data may be for a new day or that no transactions have taken place yet. The overall style suggests a focus on efficiency and ease of use relevant to the task of inventory and sales management. The visual design is professional and not overly cluttered. The use of simple icons contributes to the clean look and intuitive operation of the interface.

V. CONCLUSION

The project is a simple demonstration of how CNN can be used to solve computer vision problems with an extremely high degree of accuracy. A finger spelling sign language translator is obtained which has an accuracy of 95%. The project can be extended to other sign languages by building the corresponding dataset and training the CNN. Sign languages are spoken more in context rather than as finger spelling languages, thus, the project is able to solve a subset of the Sign Language translation problem. The main objective has been achieved, that is, the need for an interpreter has been eliminated. There are a few finer points that need to be considered when

we are running the project. The thresh needs to be monitored so that we don't get distorted grayscales in the frames. If this issue is encountered, we need to either reset the histogram or look for places with suitable lighting conditions. We could also use gloves to eliminate the problem of varying skin complexion of the signee. In this project, we could achieve accurate prediction once we started testing using a glove. The other issue that people might face is regarding their proficiency in knowing the ASL gestures. Bad gesture postures will not yield correct

prediction. This project can be enhanced in a few ways in the future, it could be built as a web or a mobile application for the users to conveniently access the project, also, the existing project only works for ASL, it can be extended to work for other native sign languages with enough dataset and training. This project implements a finger spelling translator, however, sign languages are also spoken in a contextual basis where each gesture could represent an object, verb, so, identifying this kind of a contextual signing would require a higher degree of processing and natural language processing (NLP). This is beyond the scope of this project.

VI. REFERENCES

[1] C. Sun, T. Zhang, B. K. Bao, C. Xu and T. Mei, "Discriminative exemplar coding for sign language recognition with kinect", IEEE Transactions on Cybernetics, vol. 43, no. 5, pp. 1418-1428, 2013.

[2] W. C. Hall, "What You Don't Know Can Hurt You: The Risk of Language Deprivation by Impairing Sign Language Development in Deaf Children", Maternal and Child Health Journal, pp. 1-5, 2017.

[3] R. C. Dalawis, K. D. R. Olayao, E. G. I. Ramos and M. J. C. Samonte, "KinectBased Sign Language Recognition of Static and Dynamic Hand Movements", Eighth International Conference on Graphic and Image Processing, pp. 1022501-1022501.

[4] Suharjito Ricky, Anderson Fanny, Wiryana Meita, Chandra Ariesta and Gede Putra Kusuma, "Sign Language Recognition Application Systems for Deaf-Mute People: A Review Based on Input-Process-Output", 2nd International Conference on Computer Science and Computational Intelligence 2017 ICCSCI 2017, 3-14 October 2017.

[5] A. M. Olson and L. Swabey, "Communication Access for Deaf People in Healthcare Settings: Understanding the Work of American Sign Language Interpreters", Journal for healthcare quality: official publication of the National Association for Healthcare Quality, 2016.

[6] G. Anantha Rao, K. Syamala, P.V.V. Kishore and A.S.C.S. Sastry, Deep Convolutional Neural Networks for Sign Language Recognition.

[7] Vannesa Mueller, Amanda Sepulveda and Sarai Rodriguez, The effects of baby sign training on child development In:Early Child Development and Care, vol. 184, no. 8, pp. 1178-1191, 2014.

[8] F. S. Chen, C. M. Fu and C. L. Huang, "Hand gesture recognition using a realtime tracking method and hidden Markov models", Image and vision computing, vol. 21, no. 8, pp. 745-758, 2003. Scope of the Project